

# A Mathematical Model for Fingerprinting-based Localization Algorithms

Arash Behboodi\*, Filip Lemic<sup>†</sup>, Adam Wolisz<sup>†</sup>

\*Institute for Theoretical Information Technology, RWTH Aachen University

<sup>†</sup>Telecommunication Networks Group (TKN), Technische Universität Berlin

Email: arash.behboodi@ti.rwth-aachen.de, {lemic,wolisz}@tkn.tu-berlin.de

## Abstract

Despite the popularity of Fingerprinting Localization Algorithms (FPS), general theoretical frameworks for their performance studies have rarely been discussed in the literature. In this work, after setting up an abstract model for the FPS, it is shown that fingerprinting-based localization problem can be cast as a Hypothesis Testing (HT) problem and therefore various results in the HT literature can be used to provide insights for the general FPS. This includes the scaling limits of localization reliability in terms of number of measurements and the precise characterization of a geometric error. The main quantity that encapsulates this information is shown to be the Kullback-Leibler (KL) divergence between probability distributions of a selected feature for fingerprinting at different locations. The KL divergence can be used as a central performance metric for studying the FPS, which indicates how well a localization algorithm can distinguish two points. The framework is then instantiated for Received Signal Strength (RSS)-based algorithms, where the effect of various parameters on the performance of fingerprinting algorithms is discussed, including path loss and fading characteristics, number of measurements at each point, number of anchors and their locations, and placement of training points. Simulations and experimental results characterize numerically the findings of the theoretical framework and demonstrate its consistency with realistic localization scenarios.

## Index Terms

Fingerprinting algorithms, localization, hypothesis testing, large deviation, RSS-based localization.

## I. INTRODUCTION

Precise location of people and devices, both indoors and outdoors, is an essential information for future networks as an enabler of location and context aware services, location aware and pervasive computing, and ambient intelligence. Variety of localization solutions have been studied and proposed in the recent years. These solutions exploit the spatial dependence of certain signal features. For instance, the propagation model of waves in space can be a basis for deriving certain wave features in different locations and then devising algorithms for distinguishing different locations using those features. Ultrasonic, infrared, and Radio-Frequency (RF) waves are some of the candidates for localization [1], [2], [3], [4], although the later is more popular due to its low cost and availability. RF-based solutions leverage available technologies such as IEEE 802.11 (WiFi), IEEE 802.15.4 (ZigBee), IEEE 802.15 (Bluetooth), Ultra-Wide Band (UWB), RFID, and mobile telephony. Different RF characteristics can be used for localization including Time of Arrival, Angle of Arrival, Received Signal Strength, and the quality of RF transmission in digital communication channels (e.g. Link Quality Information, Bit-Error Ratio). These features are then used as inputs in different localization algorithms. Based on signal feature processing, one can roughly distinguish three categories of localization algorithms, i.e. geometry-based, fingerprinting, and Bayesian-based ones [5]. RF fingerprinting algorithms are particularly attractive because they rely on available wireless infrastructures and therefore do not require a costly setup.

These algorithms utilize the observation that a combination of simple features of different signals is well correlated with spacial characteristic of the environment and thus rather unique at each location. In other words, the idea is to use the observed feature at a location to construct an identification tag for that location which is called the fingerprint. A database is constructed, explicitly or implicitly, by gathering fingerprints of different locations. A pattern matching algorithm identifies the location by finding the most similar fingerprint in the database to the reported fingerprint from an unknown location. The signal features should satisfy some conditions if they are to be used for reliable identification of a location. Firstly, the signal feature should be robust in the localization space. This means that the similar fingerprints should correspond to the nearby locations, otherwise a small error in fingerprinting matching can lead to large geometric error in localization. Secondly, the fingerprint should be stable in time, i.e. it should not change

significantly from the database creating time to the localization runtime, otherwise the fingerprint of the unknown location can be significantly different from the recorded fingerprint and leads to large errors. The notion of stability depends on how often the database is updated. For some cases, a fingerprint should remain stable in span of weeks and months, while, in some other cases, few seconds of stability suffices for reliable localization. The final location estimation includes similarity analysis of fingerprints and optionally post-processing methods such as k-Nearest Neighbors (kNN). An attractive feature of fingerprinting is that the analytical characterization of the relation between the signal feature and the observer's location is not required. This is usually a daunting task as it involves complex characterization of propagation and solving the wave equation for non-homogeneous environments. As a matter of fact, fingerprinting algorithms provide this functional relation between the signal feature and the location through the table of recorded samples recorded in the training database. Hence, the training database is nothing but samples of the function relating the location in space to the observed feature at the same location.

#### *A. Main Contributions*

Despite the efforts, a general theoretical framework addressing the impact of different parameters on fingerprinting performance is still missing. Most of the existing results are specific to the chosen signal feature and sometimes they do not provide the qualitative characterization for the impact of different parameters on the algorithm design. A proper theoretical formulation of fingerprinting would be useful for deriving guidelines and performance limits of fingerprinting algorithms. In this paper, we establish a connection between fingerprinting algorithms and a hypothesis testing problem. To the best of our knowledge, it is the first effort in establishing this connection. Intuitively, the signal feature is assumed to be a random variable with the probability distribution that varies with the location. In this setting, the problem of localization boils down to finding the probability distribution underlying the observed measurements, which is essentially a hypothesis testing problem. This connection leads firstly to various results regarding the performance limits of fingerprinting algorithms. Secondly, the problem of fingerprinting algorithms' design can be cast as an equivalent problem in the well-developed area of hypothesis testing. The generality of the framework makes it suitable for a variety of scenarios.

The contributions of this paper are therefore the following. We provide a general framework for theoretical study of fingerprinting algorithms using hypothesis testing problem. Based on

this framework, it is possible to define new problems related to fingerprinting algorithms such as fingerprinting under different constraints on number of training points or number of measurements. Next, we prove series of theorems establishing fundamental limits of fingerprinting algorithms. This includes a proof for existence of a fingerprinting algorithm with completely accurate localization and also showing that the inaccuracy decays exponentially with number of measurements. KL divergence between probability distributions of selected feature for fingerprinting at two different locations plays a central role in these results.

Different consequences of these theorems for tuning the parameters of fingerprinting algorithms are discussed. RSS-based algorithms are discussed next based on this framework, which are the most established and promising instances of fingerprinting algorithms. The effects of different parameters such as path loss exponent, fading statistics, and anchor selection are discussed. Finally, simulations and experimental results characterize numerically the findings of the theoretical framework and demonstrate its consistency with realistic localization scenarios.

The paper is organized as follows. Related works are discussed in Section II. In Section III, we provide a system model and a formal definition of fingerprinting algorithms. Section IV gives the theoretical limits of fingerprinting algorithms under the general system model. In Section V, we focus on a theoretical study of a RSS-based fingerprinting algorithm. Finally, the theoretical results are verified through simulation and testbed experiments in Section VI, while in Section VII we conclude the work.

## II. RELATED WORKS

There are various works in the literature studying different aspects of fingerprinting algorithms [6]. A survey of WiFi-based fingerprinting algorithms is provided in [7] and variants of such algorithms are instantiated to low-dimensional RSS [8], neural network-based clustering [9],  $K$ -means, clustering, and complexity reduction [10], and spatial signal prediction-based training algorithm [11].

The impact of the number of Access Points (APs) on the performance of fingerprinting algorithms has been studied in [12], where the authors claim that localization error is increased, when the number of APs used for constructing training database differs from number of APs used in a localization phase. In [13], the authors provide a comparison of fingerprinting algorithms based on accuracy, complexity, robustness, and scalability. The main theoretical work on fingerprinting

algorithms is [14], where the authors provide an analysis of the effect of the number of visible APs and radio propagation parameters on the performance of fingerprinting algorithms. They provide guidelines for designing and deploying a fingerprinting algorithm based on which the algorithm does not require a large number of APs. Moreover, the grid used for a training database is chosen according to the application requirements, where more dense grids provide worse accuracy but finer localization. These results are extended to complexity analysis in [15]. The authors in [16] proposed a probabilistic model of RSS-based fingerprinting relating the location to RSS. The performance of fingerprinting algorithms has then been discussed using likelihood based detection algorithms and insights have been provided for fingerprinting design. The scalability of fingerprinting algorithms is discussed in [17], where the authors suggest the scalability improvement by reducing a training database size. In [18], the authors discuss the robustness of fingerprinting to outliers and effects such as shadowing. As RSS differs across different devices, a robust fingerprinting algorithm is proposed in [19] by taking the difference of RSS of two APs as fingerprints. [20] proposes a method for estimating different antenna attenuations between different devices and by considering similarly relative differences between RSSs.

### III. SYSTEM MODEL AND DEFINITIONS

In this section, we define formally the fingerprinting algorithm using the notation given in Table I. We assume that the localization space is a connected region  $\mathcal{R}$  in  $\mathbb{R}^d$ . In practical scenarios, the dimension  $d$  is usually 2 or 3.

Symbol	Description
$\mathcal{R}$	Localization space
$\mathcal{S}$	Signal feature space
$\mathbf{S}, S$	Signal feature
$\mathcal{X}$	Fingerprint space
$\mathbf{X}, X$	Fingerprint
$\mathcal{A}$	Training grid
$\mathbf{v}$	Training points
$\mathbf{u}$	Measurement points
$\mathbf{w}$	Anchor points

TABLE I: Table of Symbols

### A. Fingerprint

Fingerprinting algorithms are based on associating a fingerprint to a location, which is later used for the identification of that location. A specific feature of environment is chosen as a basis for creating a fingerprint. We chose the term *environment* to include multiplicity of possibilities whether a particular infrastructure is assumed for localization purpose or not. The signal feature is denoted by  $S$  and belongs to the feature space  $\mathcal{S}$ . The signal is not necessarily from one source and it can be understood as combination of multiple real signals, generated by multiple sources like beacons from different APs.  $m$  consecutive observation of the signal feature  $\mathbf{S} = (S_1, \dots, S_m) \in \mathcal{S}^m$  is a random variable related to the location  $\mathbf{u}$  through the probability  $\mathbb{P}_{\mathbf{S}|\mathbf{u}}$ . In general, the consecutive measurements of signal feature can be very well dependent. Fingerprints are then constructed based on observations  $\mathbf{S}$  at different locations. Fingerprints belong to the space  $\mathcal{X}$ , which may be different from  $\mathcal{S}$ .

*Definition 1 (Fingerprint):* A fingerprint creating function  $f$  is a mapping  $\mathcal{S}^m \rightarrow \mathcal{X}^n$  that assigns to observations  $\mathbf{S}$  an element  $\mathbf{X}$  called the fingerprint at the location  $\mathbf{u}$ .

For example, if a fingerprint is the Gaussian distribution fitted to the vector of  $m$  measured received powers, then the fingerprinting space is the space of probability distribution. If one measures the Angle of Arrival (AoA) of a signal as the specific feature, then  $\mathcal{S} = [0, 2\pi)$ . If AoA is measured  $m$  times and the average value is designated as fingerprint, then  $n = 1$  and  $\mathcal{X} = [0, 2\pi)$ . However if the fingerprint is the empirical distribution of  $m$  measurements, then  $\mathcal{X}$  is the space of all probability distributions on  $[0, 2\pi)$ . Since the signal features  $\mathbf{S}$  are probabilistically related to locations, so are the fingerprints. As previously discussed, there are some requirements for a fingerprint to be useful for localization. For instance, if the fingerprint of a location is a vector of its exact distances to three fixed locations, say anchors, then the trilateration method in two-dimensional space guarantees that the fingerprint is unique to the location. In general if the probabilities  $\mathbb{P}_{\mathbf{X}|\mathbf{u}_1}$  and  $\mathbb{P}_{\mathbf{X}|\mathbf{u}_2}$  are *close enough*, i.e., if  $d(\mathbb{P}_{\mathbf{X}|\mathbf{u}_1}, \mathbb{P}_{\mathbf{X}|\mathbf{u}_2}) \leq L$  where  $d$  is a metric and  $L > 0$ . then  $\mathbf{u}_1$  and  $\mathbf{u}_2$  should also be close which is  $\|\mathbf{u}_1 - \mathbf{u}_2\| \leq s(L)$  where  $s(L) \rightarrow 0$  whenever  $L \rightarrow 0$ . We call this spatial stability of the fingerprint. The quantitative characterization of stability depends on the metric chosen for measuring *closeness* of probabilities. Any metric on the space of probability distributions can be used for this purpose. This issue will be discussed later for RSS fingerprinting.

Fingerprints should also be robust in time. This means that fingerprints should not change dramatically from the moment they are recorded until the moment they are used for identification. The robustness requirement varies with localization systems. It can be defined in terms of closeness of  $\mathbb{P}_{\mathbf{X}|\mathbf{u}_1}^{t_0}$  and  $\mathbb{P}_{\mathbf{X}|\mathbf{u}_2}^{t_1}$  where  $t_0$  and  $t_1$  are respectively training and measurement times. The notion of robustness and stability are very much related in that the spatial stability guarantees that small change in the fingerprint in time will not result in significant localization error. In this work, we assume fingerprints are robust in time.

### B. Fingerprinting Algorithm

After a signal feature and a fingerprint creating function have been selected, the next step is to design the fingerprinting algorithm. The first step is to construct a database consisting of pairs of locations and their fingerprints. Out of an uncountable set of locations in the localization space, only a finite number of locations can be chosen for direct measurement to construct the database. The training database can be constructed through extensive measurements, through simulation-based radio-map construction [21] or the combination of both. This step is called the training phase. The set of training locations, called a training grid, is denoted by  $\Lambda \subset \mathbb{R}^d$ . One can choose an algebraic structure for gridding using lattices or a non-uniform grids. The region of nearest points to each training location  $\mathbf{v} \in \Lambda$  is called its Voronoi region denoted by  $\mathcal{V}_{\mathbf{v}}$ . The training locations in  $\Lambda$  divide the localization space into regions.

*Remark 1:* Consider the Voronoi regions of nearest neighbors defined for the training locations. The closest points to a training point are not necessarily those with the closest fingerprints. One can define modified Voronoi-region  $\hat{\mathcal{V}}_{\mathbf{v}_1}$  of a training point  $\mathbf{v}_1$  as the set of all points  $\mathbf{u}$  such that their fingerprints are closer to  $\mathbf{v}_1$  than any other training points. Modified Voronoi regions are in general different from the Voronoi region.

After choosing training locations, a specific feature of environment is chosen for creating a fingerprint. The training database is created by measuring the signal feature  $\mathbf{S}$  at the training locations inside the training grid  $\Lambda$ . At each location  $\mathbf{v}$ , multiple measurements are performed and fingerprints are constructed accordingly. The training database  $\mathcal{D}$ , which consists of pairs like  $(\mathbf{v}, \mathbf{X})$ , is a subset of  $\Lambda \times \mathcal{X}^n$ .

The next part consists of finding the location of a target node. The fingerprinting algorithm reports the estimated location of the target node based on the observed signal feature. A target

node placed at the location  $\mathbf{u}$  measures the selected feature  $m'$  times and creates a fingerprint  $\mathbf{X}_{\mathbf{u}} \in \mathcal{X}^{n'}$ . In general number of measurements in the training phase and in the localization phase can be different, although we assume  $n' = n$ . After acquiring the fingerprint, a pattern matching function  $g$  is used to estimate the target node's location  $\hat{\mathbf{u}}$  based on the acquired fingerprint  $\mathbf{X}_{\mathbf{u}}$  and fingerprints in the training database. The function  $g$  is a mapping from  $\mathcal{X}^n$  to the localization space  $\mathcal{R}$ . The pattern matching function can be regarded as a composition of multiple functions. For instance a similarity kernel function can first find the most similar fingerprints in the database. Then, one can additionally use  $k$ -nearest neighbor methods to average between the  $k$  closest training locations, rather than declaring single training location. The average can be weighted or not, adaptive or non-adaptive [22]. In any case,  $k$ -NN methods will have a set of estimated locations  $\mathcal{A}_e$ , which is not equal to the original set of training locations  $\mathcal{A}$ .

One example of a similarity kernel is maximum of a function  $\Psi$ , which takes two fingerprints and calculate some a similarity score between them. In this case, the kernel first calculates the similarity score between the target node's fingerprint with the fingerprints in the fingerprinting database. Second, the estimated location is chosen as the one with the best score. The function  $\Psi$  is defined as:

$$\Psi : \mathcal{X}^n \times \mathcal{X}^n \rightarrow \mathbb{R}.$$

The location estimation is the output of the following optimization problem:

$$\begin{aligned} \hat{\mathbf{u}} &= \arg \max_{\mathbf{v} \in \mathcal{A}} \Psi(\mathbf{X}_{\mathbf{u}}, \mathbf{X}) \\ \text{subject to: } &(\mathbf{v}, \mathbf{X}) \in \mathcal{D}. \end{aligned} \tag{1}$$

In this special case, the similarity kernel is decomposed in first individual comparisons between the target node's fingerprint and the training fingerprints, followed by finding fingerprints with the maximum similarity to the target node's fingerprint. The error is obviously  $\|\mathbf{u} - \hat{\mathbf{u}}\|$ . The post-processing algorithms like kNN pick fingerprints with top  $k$  similarity scores and then the final location is offered by averaging between them. There are variants of fingerprinting algorithms which fall into our definition.

Based on the previous discussion, we can finally specify what a fingerprinting algorithm is.

*Definition 2 (Fingerprinting algorithm):* A fingerprinting algorithm for a localization space  $\mathcal{R}$  consists of:



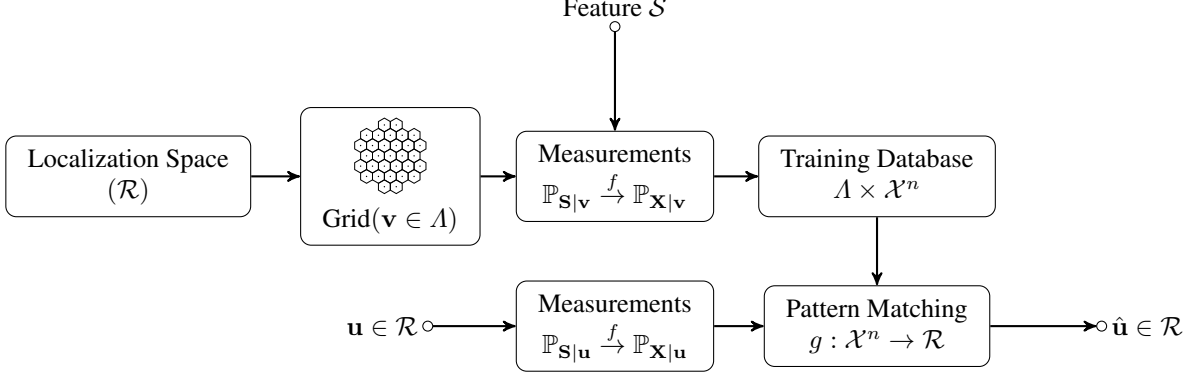


Fig. 1: Functional block of a fingerprinting algorithm

- A set of training locations  $\Lambda$ ;
- A fingerprint creating function  $f : \mathcal{S}^m \rightarrow \mathcal{X}^n$ , mapping a measured signal feature  $\mathbf{S}$  to a fingerprint  $\mathbf{X}$  in  $\mathcal{X}^n$ ;
- A training database  $\Lambda \times \mathcal{X}^n$  consisting of pairs of locations and their fingerprints;
- A pattern matching function  $g : \mathcal{X}^n \rightarrow \mathcal{R}$  reports the final location by comparing the target node's fingerprint to the ones from the training database.

Figure 1 represents the Functional block of fingerprinting algorithm starting from constructing training database to localization of an arbitrary point  $\mathbf{u}$ . The advantage of this abstraction, as it can be seen later, is its applicability in various scenarios. An example of RSS-based localization is discussed later.

### C. Performance of Fingerprinting Algorithms

After specifying fingerprinting algorithms, we introduce here a framework for evaluating the performance limit of fingerprinting algorithms. The main performance metric for any localization algorithm is the localization error. For a fingerprinting algorithm specified above, the localization error for each  $\mathbf{u} \in \mathcal{R}$  is defined as follows:

$$\Delta(\mathbf{u}) = \|\hat{\mathbf{u}} - \mathbf{u}\|.$$

Similar to the definition of error in information theory and statistics, it is possible to define the maximum and average localization error for all  $\mathbf{u}$ 's denoted by  $\Delta_{\max}$  and  $\bar{\Delta}$ . Moreover, because

we have assumed that the feature is a random variable, the localization error  $\Delta(\mathbf{u})$  can also be a random variable.

*Definition 3 (Achievable localization error):* The maximum localization error of  $\delta$  is achievable with probability  $1 - \epsilon$  if there is a fingerprinting algorithm such that :

$$\Pr(\Delta_{\max} > \delta) \leq \epsilon.$$

It is possible to similarly define achievability notion of the average localization error.

It is important to find theoretical limits of  $(\delta, \epsilon)$  pair. Next section establishes that  $(\delta, \epsilon)$  can be made arbitrarily small by increasing training points and number of measurements. Although the main goal of a fingerprinting algorithm is reliable and accurate localization, such algorithms are subject to various constraints, which introduces different trade-offs between achievable error and different fingerprinting parameters. For instance the number of training locations and the number of measurements are essentially limited and extending the infrastructure is costly. Many questions arise from this perspective. What is the minimum number of training locations  $|\mathcal{A}|$  that are required for achieving a certain localization error? How does the localization error scale with the number of measurements? How does it scale with the number of anchors? We try to address some of these questions in the paper.

#### IV. THEORETICAL LIMITS OF FINGERPRINTING ALGORITHMS PERFORMANCE

In this section, we consider the proposed general fingerprinting algorithm and find bounds on its performance. We do not assume anything about the conditional distribution  $\mathbb{P}_{X|\mathbf{u}}$  except that the measurements of the fingerprints are independent and identically distributed. Moreover it is assumed that fingerprints are spatially stable and have minimum requirements for localization.

Suppose that a target node is located either at  $\mathbf{u}_1$  or at  $\mathbf{u}_2$ . Two kinds of errors can be defined and the aim is to minimize both of these probabilities. The probabilities  $\alpha(\mathbf{u}_1, \mathbf{u}_2)$  and  $\beta(\mathbf{u}_1, \mathbf{u}_2)$  are the probability of incorrect identification.

$$\alpha(\mathbf{u}_1, \mathbf{u}_2) = \mathbb{P}(g(\mathbf{X}) = \mathbf{u}_2 | \text{Target node is at } \mathbf{u}_1) \quad (2)$$

$$\beta(\mathbf{u}_1, \mathbf{u}_2) = \mathbb{P}(g(\mathbf{X}) = \mathbf{u}_1 | \text{Target node is at } \mathbf{u}_2). \quad (3)$$

The problem of localization boils down to deciding between two probability distributions  $\mathbb{P}_{\mathbf{X}|\mathbf{u}_1}$  and  $\mathbb{P}_{\mathbf{X}|\mathbf{u}_2}$  based on the observation  $\mathbf{X}$  and subject to constraints on  $\alpha(\mathbf{u}_1, \mathbf{u}_2)$  and  $\beta(\mathbf{u}_1, \mathbf{u}_2)$ .

The problem is related to statistical hypothesis testing [23] where the goal is to decide between  $\mathbb{P}_{X|u_1}$  and  $\mathbb{P}_{X|u_2}$  based on  $n$  samples. The decision is made using a test  $\mathcal{T}$  on sample data. The errors  $\alpha(u_1, u_2)$  and  $\beta(u_1, u_2)$  are respectively missed detection and false alarm. This connection enable us to regard a design problem in fingerprinting localization as a hypothesis testing problem and benefit from abundant research materials in that area. As a first step, the following theorem provides fundamental limits on the performance of fingerprinting algorithms.

*Theorem 4.1:* Consider a fingerprinting algorithm with fingerprints  $\mathbf{X} \in \mathcal{X}^n$  consisting of  $n$  i.i.d. fingerprints from conditional distribution as  $\mathbb{P}_{X|u}$ . If the distribution  $\mathbb{P}_{X|u}$  is known to the fingerprinting algorithm, then there is a pattern matching function  $g$ , such that for each pair of locations  $(u_1, u_2)$  and for any  $0 < \epsilon < 1$ ,

$$\alpha(u_1, u_2) \leq \epsilon \quad (4)$$

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log \beta(u_1, u_2) = -D(\mathbb{P}_{X|u_1} \parallel \mathbb{P}_{X|u_2}). \quad (5)$$

*Proof:* The proof of this theorem as well as other theorems in this section can be found in large deviation theory or hypothesis testing texts. Nevertheless we prove this theorem in Appendix to point out the idea behind these theorems. The proof mainly consists of finding a set that contains the observations with high probability. This *typical* set then serves a decision region for declaring a hypothesis as true. ■

Theorem 4.1 provides fundamental bounds on error probability if maximum localization error is intended. In other words, a fingerprinting algorithm achieves the best error probabilities subject to localizing every location in the localization space. The theorem shows that if  $D(\mathbb{P}_{X|u_1} \parallel \mathbb{P}_{X|u_2})$  is non-zero, i.e. strictly positive, then the probability of error asymptotically decays exponentially with  $D(\mathbb{P}_{X|u_1} \parallel \mathbb{P}_{X|u_2})$ . It can also be shown that this is the best exponent one can get under the first constraint on  $\alpha(u_1, u_2)$  [23]. If the KL-divergence of two probability distributions is small, this means that the error decays with smaller exponent and therefore more measurements are needed to guarantee a fast decay in error. In other words,  $n$  should be increased to compensate for small divergence and reduce the error. More observations are needed for identifying the difference between very similar phenomenons.

One can also infer the optimal fingerprint construction using insights from hypothesis testing problem. From HT point of view, the optimal fingerprint is the sufficient statistics and the pattern

matching function is the statistical test. It is known that Neyman-Pearson tests are optimal tests in the sense that keeping one of the error fixed, they achieve the minimum possible value for the other error. A Neyman-Pearson test is a test in which the normalized observed log-likelihood ratio is compared with a threshold  $\gamma$ . In other words, based on the observations  $\mathbf{X}$ , the normalized log-likelihood ratio  $T_n$  is defined as:

$$T_n = \frac{1}{n} \log \frac{\mathbb{P}_{\mathbf{X}|\mathbf{u}_1}}{\mathbb{P}_{\mathbf{X}|\mathbf{u}_2}}. \quad (6)$$

If  $T_n$  is bigger than  $\gamma$ ,  $\mathbf{u}_1$  is announced as the location and otherwise,  $\mathbf{u}_2$  is announced.

For Neyman-Pearson test, there are lot of results characterizing the asymptotic and non-asymptotic behavior of the errors in (4) and (5). One particularly relevant parallel can be established when the target node is present at each location with an a priori probability. This is an example of map-aware localization. Knowing the map means that one knows a priori the probability that a target node is present at each location. This is equivalent to finding the best Bayes probability of error defined as:

$$P_n^{(e)} = \mathbb{P}(\mathbf{u}_1)\alpha(\mathbf{u}_1, \mathbf{u}_2) + \mathbb{P}(\mathbf{u}_2)\beta(\mathbf{u}_1, \mathbf{u}_2).$$

Interestingly Neyman-Pearson test with  $\gamma = 0$  is the optimal test.

*Proposition 1:* If the probability that the target node is present at  $\mathbf{u}_1$  is in  $(0, 1)$ , then there is a map-aware fingerprinting algorithm such that:

$$\liminf_{n \rightarrow \infty} \frac{1}{n} \log P_n^{(e)} = -I_0(0), \quad (7)$$

where  $I_0(0)$  is called the Chernoff information of the probabilities  $\mathbb{P}_{\mathbf{X}|\mathbf{u}_1}$  and  $\mathbb{P}_{\mathbf{X}|\mathbf{u}_2}$  and is defined as:

$$I_0(0) = -\log \inf_{t \in \mathbb{R}} \mathbb{E} \left( \exp \left( t \log \frac{\mathbb{P}_{\mathbf{X}|\mathbf{u}_2}}{\mathbb{P}_{\mathbf{X}|\mathbf{u}_1}} \right) \right)$$

and the expectation is with respect to  $\mathbb{P}_{\mathbf{X}|\mathbf{u}_1}$ .

The proof is exact replication of equivalent hypothesis testing problem and follows from large deviation theory analysis. It can be found in [24]. The important insight is that map-aware localization can be done using  $T_n$  in (6) with the threshold zero and the error decays exponentially with  $n$  and with the Chernoff information  $I_0(0)$ . Many other similar and interesting conclusions can be made from the hypothesis testing analogy. For instance, if Neyman-Pearson test is used, the errors decay exponentially with rates dependent on the choice of threshold  $\gamma$ , similar to Proposition 1.

The errors in (4) and (5) are obtained under the strong assumption that an algorithm knows the conditional probability distribution. One point of using fingerprinting is exactly to avoid the complexity behind characterizing the relation between locations and signal features. Hence, it is interesting to understand what happens if the conditional probability distribution is not known. One option is to use the training phase to learn the probability distribution at each location. Suppose that fingerprints take their value on a finite space  $\mathcal{X}$ . The assumption of finiteness is natural since the measurements are usually quantized and scaled version of the signal feature. Suppose that in the training phase,  $\mathbf{X}$  is observed at  $\mathbf{u}$ . The empirical distribution of  $\mathbf{X}$  can be defined as:

$$\mathbb{Q}_{\mathbf{X}|\mathbf{u}}(x) = \frac{1}{n} \sum_{i=1}^n \mathbf{1}(x = X_i).$$

For a finite fingerprinting space  $\mathcal{X}$ , from the law of large numbers it follows that  $\mathbb{Q}_{\mathbf{X}|\mathbf{u}}$  will be almost surely equal to  $\mathbb{P}_{\mathbf{X}|\mathbf{u}}$  as  $n$  tends to infinity. Therefore, with large  $n$  it is guaranteed that the conditional probability distribution  $\mathbb{P}_{\mathbf{X}|\mathbf{u}}$  is recovered during the training phase.  $\mathbb{Q}_{\mathbf{X}|\mathbf{u}}(x)$  is sum of  $n$  independent random variables  $\mathbf{1}(x = X_i)$  with support in  $\{0, 1\}$ . Using Hoeffding's inequality [25], it can be seen that the empirical distribution converges point-wise to the true distribution exponentially fast with the number of measurements. This writes as:

$$\Pr(|\mathbb{Q}_{\mathbf{X}|\mathbf{u}}(x) - \mathbb{P}_{\mathbf{X}|\mathbf{u}}(x)| \geq a) \leq 2e^{-2na^2}. \quad (8)$$

Note that the decay exponent is  $a^2$ . It is possible to derive the exact exponent using large deviation theory. To this end, the total variation distance is used as a measure for closeness of probabilities. It is defined as follows:

$$d(\mathbb{Q}_{\mathbf{X}|\mathbf{u}}, \mathbb{P}_{\mathbf{X}|\mathbf{u}}) = \frac{1}{2} \sum_{x \in \mathcal{X}} |\mathbb{Q}_{\mathbf{X}|\mathbf{u}}(x) - \mathbb{P}_{\mathbf{X}|\mathbf{u}}(x)|.$$

Law of large numbers say that  $d(\mathbb{Q}_{\mathbf{X}|\mathbf{u}}, \mathbb{P}_{\mathbf{X}|\mathbf{u}})$  tends to zero as  $n$  tends to infinity. Total variation distance and KL-divergence are related through Pinsker's inequality [26] as follows:

$$d(\mathbb{Q}, \mathbb{P}) \leq \sqrt{\frac{1}{2} D(\mathbb{Q} || \mathbb{P})}.$$

The inequality is one-sided, since there is in general probabilities with zero total variation distance and infinite KL-divergence. The following proposition provides bounds on how fast the empirical distribution tends to the original distribution.

*Proposition 2:* A fingerprinting localization algorithm can estimate  $\mathbb{P}_{X|\mathbf{u}}$  by the empirical distribution  $\mathbb{Q}_{X|\mathbf{u}}$  calculated based on  $n$  i.i.d. samples  $X_i$  at  $\mathbf{u}$  and it follows:

$$\begin{aligned} \lim_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P}(d(\mathbb{Q}_{X|\mathbf{u}}, \mathbb{P}_{X|\mathbf{u}}) > a) \\ = - \inf_{\mathbb{P} \in B(a, \mathcal{X})} D(\mathbb{P} || \mathbb{P}_{X|\mathbf{u}}) \end{aligned} \quad (9)$$

where

$$B(a, \mathcal{X}) = \{\mathbb{P} : d(\mathbb{P}, \mathbb{P}_{X|\mathbf{u}}) > a\}.$$

*Proof:* The proposition is direct consequence of Sanov's theorem [24]. Note that a similar analysis can be conducted for a general feature space  $\mathcal{X}$ , not necessarily finite. We avoid the technicalities here as the main conclusions remain similar. For the general analysis refer to [24, Chapter 6.]. ■

The previous proposition guarantees that the probability distribution can be estimated well during the training phase using  $n$  number of measurements. Moreover the empirical distribution  $\mathbb{Q}_{X|\mathbf{v}}$  can act as the fingerprint. Using Pinsker's inequality, it can be seen that  $D(\mathbb{P} || \mathbb{P}_{X|\mathbf{u}}) \geq 2a^2$  for  $\mathbb{P} \in B(a, \mathcal{X})$  which recovers the exponent of Hoeffding's inequality.

Hence, we have shown that a fingerprinting algorithm does not need to know a priori the probability distribution relating a feature to a location. In general, for a spatially stable fingerprint, there is a fingerprinting algorithm that can detect all locations without knowing the conditional probability distributions from enough large number of measurements. The error probabilities decay exponentially with the number of measurements. Similar conclusions hold if the finiteness of  $\mathcal{X}$  is not assumed but it requires some technicalities that we avoid here.

#### A. Fingerprinting with Training Grids

So far, a two-point localization scenario has been considered. This can be easily extended to a localization scenario with finite candidate locations. Basically FPS can reliably and accurately localize a target node within finite possible locations using enough large number of measurements. When possible locations are uncountable, the localization space is divided into multiple regions and the equivalent HT problem aims at finding in which region the target node is located. These regions are indeed related to the training grid. No assumption was made regarding the training database in previous part. Indeed, in order to have perfectly accurate fingerprinting, training

fingerprints are required for every points which is practically impossible. Suppose that training locations are limited to those in  $\Lambda$ , which is a finite set. The fingerprinting algorithm announces, similar to (1), a training point as the estimated location. If a user is located at  $\mathbf{u}$ , then the question is how well one can localize it by comparison to the training database. Following the discussion in Remark 1, the estimated location is not necessarily the closest training point to the location  $\mathbf{u}$  geometrically. Based on the chosen fingerprint and the pattern matching function, the points that are estimated by a training point create the modified Voronoi region of that point denoted by  $\hat{\mathcal{V}}_{\mathbf{v}}$ . Localization in this scenario is equivalent to finding the modified Voronoi region  $\hat{\mathcal{V}}_{\mathbf{v}}$  that contains the target node, thus determines the closest training location to the node. Similarly, the following errors can be defined for this scenario:

$$\alpha(\hat{\mathcal{V}}_{\mathbf{v}_1}, \hat{\mathcal{V}}_{\mathbf{v}_2}) = \mathbb{P}(g(\mathbf{X}) = \mathbf{v}_2 \mid \text{Target node is inside } \hat{\mathcal{V}}_{\mathbf{v}_1}).$$

$$\beta(\hat{\mathcal{V}}_{\mathbf{v}_1}, \hat{\mathcal{V}}_{\mathbf{v}_2}) = \mathbb{P}(g(\mathbf{X}) = \mathbf{v}_1 \mid \text{Target node is inside } \hat{\mathcal{V}}_{\mathbf{v}_2}).$$

Following theorem provides the fundamental limit of fingerprinting localization using finite number of training location.

*Theorem 4.2:* Consider a fingerprinting algorithm with training locations in  $\Lambda$ . For the fingerprint  $\mathbf{X}$ , there is a pattern matching function  $g$ , based on the empirical distribution, such that for each  $\epsilon > 0$  and enough large  $n$ ,  $\alpha(\hat{\mathcal{V}}_{\mathbf{v}_1}, \hat{\mathcal{V}}_{\mathbf{v}_2}) \leq \epsilon$  and for  $\mathbf{u}_2 \notin \hat{\mathcal{V}}_{\mathbf{v}_1}$ , we have:

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log \beta(\hat{\mathcal{V}}_{\mathbf{v}_1}, \hat{\mathcal{V}}_{\mathbf{v}_2}) = -D(\mathbb{P}_{X|\hat{\mathcal{V}}_{\mathbf{v}_1}} \parallel \mathbb{P}_{X|\mathbf{u}_2}), \quad (10)$$

where :

$$D(\mathbb{P}_{X|\hat{\mathcal{V}}_{\mathbf{v}_1}} \parallel \mathbb{P}_{X|\mathbf{u}_2}) = \inf_{\mathbf{u} \in \hat{\mathcal{V}}_{\mathbf{v}_1}} D(\mathbb{P}_{X|\mathbf{u}} \parallel \mathbb{P}_{X|\mathbf{u}_2}).$$

*Proof:* See Appendix. ■

Theorem 4.2 guarantees that fingerprinting algorithms can successfully find in which region the target node is located. The region is the modified Voronoi region. The maximum error is obtained for those that are closer to the boundary of the region. For the locations exactly on the boundary, the algorithm might report multiple training points as estimation. An application of  $k$ -nearest neighbor in this situation can reduce the error significantly.

### B. On Training Grid Selection

From Theorem 4.2, the importance of spatial stability of fingerprints becomes more clear. Without stability, the modified Voronoi region can be so different from the original Voronoi regions that the closest fingerprint in the training database corresponds to a training point that is very far from the target node and hence leading to huge error. The notion of covering radius, borrowed from lattice literature [27], can be useful here. The covering radius of a region  $\mathcal{V}_{\mathbf{v}}$  is defined as follows:

$$r_{\text{cov}}(\mathcal{V}_{\mathbf{v}}) = \inf\{r : \mathcal{V}_{\mathbf{v}} \subseteq \mathbf{v} + r\mathcal{B}\}$$

where  $\mathcal{B}$  is unit-radius ball in space. The covering radius is indeed the smallest ball centering at  $\mathbf{v}$  and covering the region. The covering radius will provide the maximum localization error for the algorithm of Theorem 4.2.

*Corollary 1:* Consider a fingerprinting algorithm with training locations in  $\Lambda$  and modified Voronoi regions  $\hat{\mathcal{V}}_{\mathbf{v}}$ . The maximum error  $\Delta_{\max}$  is equal to the maximum of all covering radius for each modified Voronoi regions:

$$\Delta_{\max} = \max_{\mathbf{v} \in \Lambda} r_{\text{cov}}(\hat{\mathcal{V}}_{\mathbf{v}}).$$

Stability of fingerprints guarantees that the covering radius of modified Voronoi regions does not change that much from the original Voronoi region. Throughout this section, we have assumed the spatial stability of fingerprints. However, if the KL-divergence  $D(\mathbb{P}_{X|\mathcal{V}_{\mathbf{v}_1}} \|\mathbb{P}_{X|\mathcal{V}_{\mathbf{v}_2}})$  for two training locations is zero then one cannot distinguish between them using a sample feature. This will contribute to the maximum error. This can specially happen when very fine training grids is considered. As the granularity of a training grid is increased, the Voronoi region becomes smaller but their modified Voronoi regions can hugely overlap. This means that the covering radius of Voronoi region is decreased but the covering radius of modified Voronoi region remains the same. The conclusion is that using finer grids is sometimes beyond need since it does not increase the accuracy.

Localization error depends also on the shape of Voronoi regions and modified Voronoi regions. Here we assume that those regions are the same. The choice of training locations affects the covering radius and the maximum distance and whereby affects performance of fingerprinting algorithms. For example, the hexagonal lattice with the minimum distance between locations  $d$



has the covering radius of  $\frac{d}{\sqrt{3}}$ . The square lattice with the same minimum distance  $d$  has larger covering radius  $\frac{d}{\sqrt{2}}$ . This means that the hexagonal grid is better than square lattice with respect to its coverage radius while keeping same minimum distance between training locations.

## V. RSS-BASED FINGERPRINTING ALGORITHMS

So far, the feature used for fingerprinting has not been explicitly specified and it is modeled as a general random variable. The most common choice for the feature is RSS from multiple anchors. Multiple anchors belong to a wireless network which is also used for fingerprinting localization. The anchors are  $K$  APs regularly transmitting their signals. The anchors are not interfering with each other. This can be obtained by employing proper Media Access Control (MAC) mechanisms, although we do not consider specific MAC mechanisms in this work. We assume that the anchors are placed in a 2-dimensional Euclidean space. The anchor  $i$  is placed at the location  $\mathbf{w}_i \in \mathbb{R}^2$  and transmits with the transmission power  $P_T^{(i)}$ . The transmitted signal in time is presented as  $x^{(i)}(t)$  and RSS at the location  $\mathbf{u}$  as  $y^{(i)}(\mathbf{u}, t)$ . The relation between  $x^{(i)}(t)$  and  $y^{(i)}(\mathbf{u}, t)$  is explicated in the next part.

### A. Channel Model

In order to explicitly express the dependence of RSS from an anchor on the distance, the received signal is modeled as follows [28]:

$$y^{(i)}(\mathbf{u}, t) = \sum a_j^{(i)}(t) x^{(i)}(t - \tau_j^{(i)}(t)) + z^{(i)}(t)$$

where  $a_j^{(i)}(t)$  and  $\tau_j^{(i)}(t)$  are the channel gain and delay of  $j$ 'th multi-path component. The variation of these parameters determine the type of environment, namely indoor and outdoor, fast fading or slowly fading, frequency selective or frequency flat. One option is to consider statistical channel model and take  $a_j^{(i)}[n] = a_j^{(i)}(nT)$  as Rayleigh distributed ( $T$  is sampling period). For simplicity, we consider only one dominant channel tap, which is  $a_1^{(i)}[n]$  and  $a_j^{(i)}[n] = 0$  for  $j > 1$ .  $a_1^{(i)}[n]$ 's are correlated for different  $n$  and their variation depends on the coherence time of the channel. They also incorporate the path loss. Two main scenarios are considered in this work.

First it is assumed that the power is calculated over long period and therefore they are stable in time. Therefore  $a_1^{(i)}[n]$  is assumed to be  $\frac{1}{\|\mathbf{w}_i - \mathbf{u}\|^{\alpha_i/2}}$ . For example, WiFi RSSI values vary

slightly in time, mainly due to quantization noise. Using this model, the fingerprint  $X_{\mathbf{u}}^{(i)}$  is the received power and it is obtained as:

$$X_{\mathbf{u}}^{(i)} = P^{(i)}(\mathbf{u}) = \frac{P_T^{(i)}}{\|\mathbf{w}_i - \mathbf{u}\|^\alpha} + N + \mathbf{N}_i, \quad (11)$$

where  $\alpha$  is the path loss exponent,  $N$  is the additive noise power,  $\mathbf{N}_i$  is a Gaussian random variable of variance  $N_i$  to account for small changes in RSS values. The fingerprint at the point  $\mathbf{u}$  is then  $\mathbf{X}_{\mathbf{u}} = (X_{\mathbf{u}}^{(1)}, \dots, X_{\mathbf{u}}^{(K)})$ .

In second scenario it is assumed that channel coefficient changes in each round of RSS calculation. This means that  $a_1^{(i)}[n]$  is assumed to be  $\frac{\sqrt{\mathbf{H}^{(i)}}}{\|\mathbf{w}_i - \mathbf{u}\|^{\alpha_i/2}}$  where  $\mathbf{H}^{(i)}$  is a random variable representing multipath fading and shadowing effect. No additional quantization noise is assumed. Fingerprint is chosen as RSS value and it is calculated as follows:

$$X_{\mathbf{u}}^{(i)} = P^{(i)}(\mathbf{u}) = \frac{\mathbf{H}^{(i)} P_T^{(i)}}{\|\mathbf{w}_i - \mathbf{u}\|^{\alpha_i}} + N. \quad (12)$$

A database is created by measuring RSS at the training locations  $\mathbf{v} \in \mathcal{A}$ . The measurements are used for localization of a target node located at the location  $\mathbf{u} \in \mathcal{R}$ . At each training location  $\mathbf{v}$ , RSS is measured from anchor  $i$  and the measurement is repeated for number of times.

In the next parts, we study these scenarios using the framework developed above. Based on the results of previous section, the KL-divergence  $D(\mathbb{P}_{\mathbf{X}|\mathbf{u}_1} \|\mathbb{P}_{\mathbf{X}|\mathbf{u}_2})$  is adopted as the main metric of interest. This metric is inversely related with the latency of localization. Bigger KL-divergence indicates fewer measurements required for localization.

### B. Noisy RSS Fingerprinting

For the choice of fingerprints (11), the metric  $D(\mathbb{P}_{\mathbf{X}|\mathbf{u}_1} \|\mathbb{P}_{\mathbf{X}|\mathbf{u}_2})$  is KL divergence of two normal random variable and it is evaluated as:

$$D(\mathbb{P}_{\mathbf{X}|\mathbf{u}_1} \|\mathbb{P}_{\mathbf{X}|\mathbf{u}_2}) = \sum_j \frac{(P_T^{(j)})^2}{2N_j} \left( \frac{1}{\|\mathbf{w}_j - \mathbf{u}_1\|^\alpha} - \frac{1}{\|\mathbf{w}_j - \mathbf{u}_2\|^\alpha} \right)^2. \quad (13)$$

This metric corresponds well to our geometric intuition since if  $D(\mathbb{P}_{\mathbf{X}|\mathbf{u}_1} \|\mathbb{P}_{\mathbf{X}|\mathbf{u}_2})$  is zero, then  $\|\mathbf{w}_j - \mathbf{u}_2\| = \|\mathbf{w}_j - \mathbf{u}_1\|$  for all anchors. Therefore to guarantee uniqueness, the anchors should be placed in a way to avoid this situation which can be achieved simply by choosing three non-collinear anchors in  $\mathbb{R}^2$ . As it will be discussed later, this metric can provide more guidelines

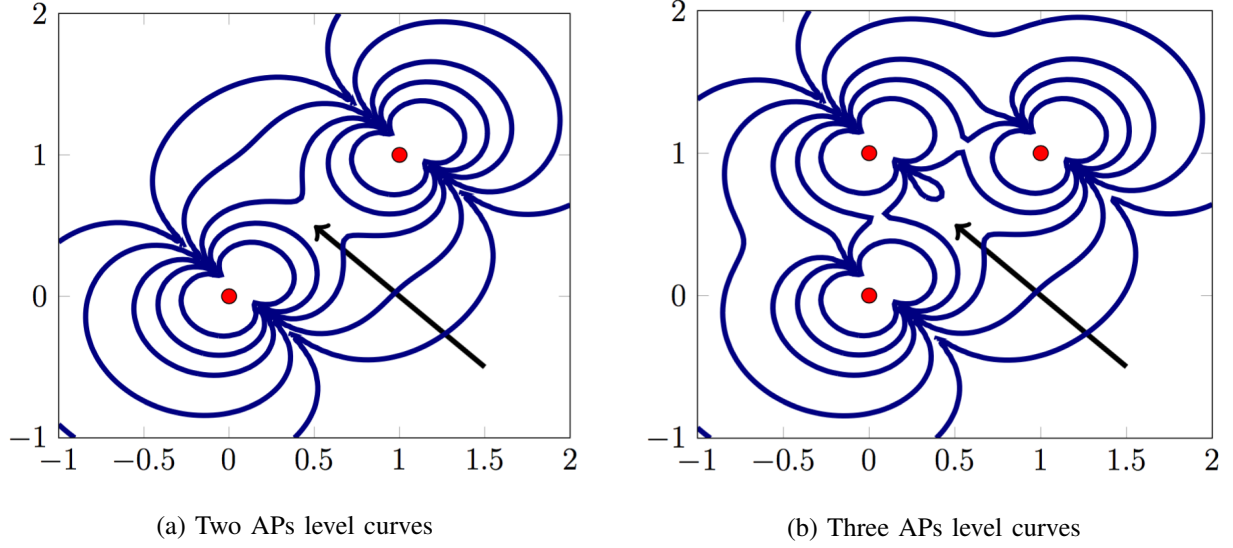


Fig. 2: Level curves of KL-divergence (localization latency)

for anchor placement. To understand stability of the fingerprint, we consider localization over  $\mathbb{R}^+$  with an anchor placed at origin and the localization space inside  $[0, D]$ .

*Proposition 3:* Suppose that the localization space lies in  $[0, D]$ . For an anchor placed at origin, RSS-fingerprints are spatially stable. In other words, if  $D(\mathbb{P}_{\mathbf{X}|u_1} \|\mathbb{P}_{\mathbf{X}|u_2}) \leq L$  then:

$$|u_1 - u_2| \leq \frac{D^{\alpha+1}}{\alpha P_T} \sqrt{2N_1 L}$$

*Proof:* See Appendix. ■

Note that if the locations are arbitrarily far from the anchor, then  $|u_1 - u_2|$  can be arbitrarily large even for small  $L$ . Therefore RSS fingerprints are not stable in general for an unbounded localization space. The proposition 3 indicates also how far one can measure RSS of an anchor to guarantee a certain accuracy. Moreover the stability is improved with the transmission power of the anchor and path loss exponent. The relation between  $|u_1 - u_2|$  and  $D(\mathbb{P}_{\mathbf{X}|u_1} \|\mathbb{P}_{\mathbf{X}|u_2})$  can be used to provide bounds on the accuracy. For instance consider a typical setting for WiFi fingerprinting. Let the localization region be a line of 10 meter, i.e.,  $D = 10m$ , transmission power be  $P = 100mW$ , RSS noise be around  $N_1 = 10^{-3}$  and the granularity of devices permit separating fingerprints by precision of  $L = 10^{-3}$ . Those points that have  $D(\mathbb{P}_{\mathbf{X}|u_1} \|\mathbb{P}_{\mathbf{X}|u_2}) \leq L = 0.001$  are at most 5 cm apart, which means that the FPS can achieve at best around 5 cm geometric error. This provides guidelines on how the training locations should be chosen to

guarantee a certain accuracy. Note that the stability is improved with the transmission power of the anchor.

Consider again the metric  $D(\mathbb{P}_{\mathbf{X}|\mathbf{u}_1} \|\mathbb{P}_{\mathbf{X}|\mathbf{u}_2})$  in general. From this, some insights can be directly derived. First of all, increasing number of anchors can improve KL-divergence and thereby reduce the number of required measurements for localization. This is subject to proper placement of anchor points. Note that the worst performance are obtained at the points with smallest KL-divergence and therefore new anchors should be placed in a way to increase the KL-divergence exactly for those points. To see this, suppose that the goal is to be able to distinguish all points with distance more than  $d$ . Consider all points  $\mathbf{u}_2$  and  $\mathbf{u}_1$  such that  $\|\mathbf{u}_2 - \mathbf{u}_1\| = d$ . One way to study this problem is to look at level curves of the function:

$$\ell(\mathbf{u}, \mathbf{e}) = \sum_j \frac{(P_T^{(j)})^2}{2N_j} \left( \frac{1}{\|\mathbf{w}_j - \mathbf{u}\|^\alpha} - \frac{1}{\|\mathbf{w}_j - \mathbf{u} - \mathbf{e}\|^\alpha} \right)^2,$$

where  $\mathbf{e}$  is an arbitrary vector of norm  $d$ . Figure 2 shows these level curves for two and three anchors and  $\mathbf{e} = (0.1, 0.1)$  where the arrows show the direction of increase in KL-divergence. Changing  $\mathbf{e} = (0.1, 0.1)$  rotates the plots according to the new chosen vector. This does not change significantly the shape of curves far from anchors unlike those close to anchors. If the number of measurements to achieve certain accuracy is an indicator of latency, these curves show the latency of localization for distinguishing  $\mathbf{u}$  and  $\mathbf{u} + \mathbf{e}$ .

It can be seen that the location pairs with the same  $\|\mathbf{u}_2 - \mathbf{u}_1\|$  that are closer to APs have larger difference in their fingerprints. The larger difference between fingerprints provide more stable scheme to eventual localization errors.

It can be seen that the small value of  $\ell(\mathbf{u}, \mathbf{e})$ , which indicates bad localization performance, corresponds to an oval surrounding the anchors and particularly to points between the anchors. This observation suggests that a new anchor should be placed on those curves containing the localization area. Moreover one option to find the position of new anchors is to consider Voronoi regions of current anchors and place new anchor on the intersection of Voronoi regions. In this way, KL-divergence is increased by installing nearby anchors to compensate the effect of far anchors. An analytical explanation of this phenomenon can be given by using mean value theorem for several variables. Namely it can be seen that:

$$D(\mathbb{P}_{\mathbf{X}|\mathbf{u}_1} \|\mathbb{P}_{\mathbf{X}|\mathbf{u}_2}) \leq \sum_j \frac{\alpha^2 (P_T^{(j)})^2}{2N_j} \left( \frac{\|\mathbf{u}_1 - \mathbf{u}_2\|^2}{\|\bar{\mathbf{u}}_j - \mathbf{w}_j\|^{2\alpha+2}} \right)$$

where  $\bar{\mathbf{u}}_j$  is a location on the line between  $\mathbf{u}_1$  and  $\mathbf{u}_2$ . For enough far points from all anchors, all  $\bar{\mathbf{u}}_j$  can be approximated as equal and one can use the function  $\tilde{\ell}(\mathbf{u}) = \sum_j \frac{1}{\|\mathbf{u} - \mathbf{w}_j\|^{2\alpha+2}}$  for evaluating the effect of anchor placements on localization performance. First, for enough small or enough large path loss exponent  $\alpha$ , the distance between two fingerprints is arbitrarily small. Moreover the level curves of  $\tilde{\ell}(\cdot)$  approximate well that of  $\ell(\cdot)$  for far points from anchors.  $\tilde{\ell}$  provides a better understanding in general since it does not depend on  $\mathbf{e}$ .

### C. Fingerprinting under Fading

In this part, we consider the situation in which we assume the channel gains are random variables. The randomness can be due to reasons including shadowing and multi-path fading. For the anchor  $i$ , the RSS value is modeled as:

$$X_{\mathbf{u}}^{(i)} = \frac{\mathbf{H}^{(i)} P_T^{(i)}}{\|\mathbf{w}_i - \mathbf{u}\|^{\alpha_i}} + N.$$

Let us again assume a single anchor scenario. From Theorem 4.1, we look at the KL-divergence of RSS values for different locations. Considering standard Rayleigh distribution for fading,  $\mathbf{H}^{(i)}$  is exponentially distributed and we have:

$$D(\mathbb{P}_{\mathbf{X}|\mathbf{u}_1} \|\mathbb{P}_{\mathbf{X}|\mathbf{u}_2}) = \sum_j \left( \log \frac{\|\mathbf{w}_j - \mathbf{u}_2\|^\alpha}{\|\mathbf{w}_j - \mathbf{u}_1\|^\alpha} + \frac{\|\mathbf{w}_j - \mathbf{u}_1\|^\alpha}{\|\mathbf{w}_j - \mathbf{u}_2\|^\alpha} - 1 \right).$$

To compare this with noisy case, a standard inequality  $x - 1 - \log x \geq \frac{1}{2}(x - 1)^2$  for  $x < 1$  is used to show that:

$$\log \frac{\|\mathbf{w}_j - \mathbf{u}_2\|^\alpha}{\|\mathbf{w}_j - \mathbf{u}_1\|^\alpha} + \frac{\|\mathbf{w}_j - \mathbf{u}_1\|^\alpha}{\|\mathbf{w}_j - \mathbf{u}_2\|^\alpha} - 1 \geq \frac{\|\mathbf{w}_j - \mathbf{u}_1\|^{2\alpha}}{2} \left( \frac{1}{\|\mathbf{w}_j - \mathbf{u}_1\|^\alpha} - \frac{1}{\|\mathbf{w}_j - \mathbf{u}_2\|^\alpha} \right)^2. \quad (14)$$

Without loss of generality, we assumed that  $\|\mathbf{w}_j - \mathbf{u}_1\| \leq \|\mathbf{w}_j - \mathbf{u}_2\|$ . The right hand side of the inequality in (14) shows an additional product factor  $\|\mathbf{w}_j - \mathbf{u}_1\|^{2\alpha}$  compared to (13) ignoring the constant factors. The inequality in (14) shows that for points far from the anchors ( $\|\mathbf{w}_j - \mathbf{u}_1\| > 1$ ), KL-divergence is larger in fading environments and therefore localization performance is superior. The situation is reverse for points closer to the anchors,  $\|\mathbf{w}_j - \mathbf{u}_1\| < 1$ . This can also be verified using level curves of  $\ell(\mathbf{u}, \mathbf{e})$  as defined above in Figure 3. In other words, shadowing and fading are expected to improve the worse case performance of fingerprinting algorithms intuitively because they create more variability in the RSS pattern of different locations. Regarding anchor placement, level curves of  $\ell(\mathbf{u}, \mathbf{e})$  suggest that the same guideline regarding Voronoi-based anchor placement holds far fading case too.

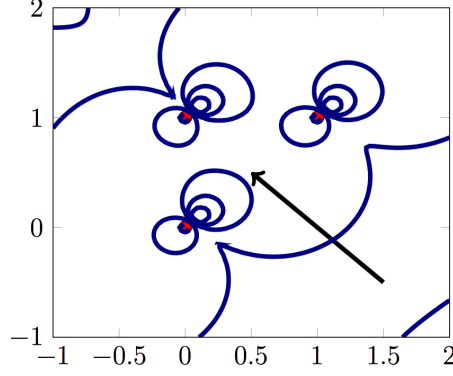


Fig. 3: Level curves of KL-divergence (localization latency) under fading

## VI. SIMULATION AND EXPERIMENTAL RESULTS

In this section, numerical simulations are used to characterize quantitatively the findings of the proposed theory within a model close to our theoretical assumptions and yet more complex with a finite number of training locations and measurements per location. Our experimental setup is then employed to show that our theoretical guidelines still hold in a realistic environment for the same algorithm. As a fingerprint at each location, we select the vector of average RSS values observed from different WiFi APs, which is a well-known fingerprint selection method [22], [29]. A similarity kernel is the Euclidean distance between vectors, which is again a standard method used in fingerprinting [30].

### A. Simulation Results

In our simulation environment, we define a set of APs related parameters, i.e. their locations and transmit powers. RSS values obtained from each AP at a targeted node with an unknown location are modeled using the COST 231 multi-wall model for indoor radio propagation [31]. The applicability of the model has been demonstrated for localization purposes [32] and the model has already been extensively used (e.g. [33]). The model accounts the type and number of walls, floors or obstacles in an environment, as well as the locations of APs. The first attenuation contribution in the model is a well-known one-slope term that relates received power to distance. Two parameters influence the attenuation in this term: the constant  $l_0$  (the path-loss at 1 m distance and at the center frequency of 2.45 GHz) and the path-loss exponent  $\gamma$ . The second attenuation contribution is a linear wall attenuation term. The number of walls in the direct path between an AP and a target node is counted for each wall an attenuation contribution is assumed. The model outputs RSS values from the defined APs at a target node's location. A noise is then

introduced to the derived RSS values, where the noise is drawn from a Gaussian distribution  $\mathcal{N}(0, \sigma)$ . Introduction of a Gaussian modeled noise is an extensively used approach in indoor positioning related simulations [34].

For the simulation and later experimental examination environment we selected the TWIST testbed [35]. The TWIST testbed environment is an office building, with its outline given in Figure 10. In the model parameterization, we used measurements from the TWIST testbed and leveraged a least square fitting procedure that allows minimizing the cost function between the measured received power and the modeled one. The parameters that are given to the model are the constant  $l_c$  related to the least square fitting procedure, the path-loss exponent  $\gamma$ , and the wall attenuation factor  $l_w$ . Additionally, a zero-mean Gaussian noise with standard deviation  $\sigma$  has been added to the obtained RSS values. If not explicitly stated otherwise, for deriving our simulation results we used  $l_c = 53.73$ ,  $\gamma = 1.64$ ,  $l_w = 4.51$ ,  $\sigma = 2$ . The transmit power of each AP equals 20 dBm. In our simulation, we defined a set of 4 AP, with their locations indicated in Figure 10. A target's node location has been selected randomly, its location has been estimated using the selected fingerprinting algorithm, and the localization error, i.e. an offset from the true location has been calculated. The procedure has been repeated 10000 times and the results have been reported in a regular box-plot fashion.

First, we evaluate the statement given in Theorem 4.2 in regards to the density of training locations and the selection of the training grid. We compare the performance of the algorithm in case a selected training grid is hexagonal, squared and random, as shown in Figure 4. We further evaluate the effect of the number of points in a training grid, hence in addition to the depicted grids that each contains 40 training locations (3 m cell size for the squared grid), we also evaluate the case with 105 (2 m cell size) and 420 (1 m cell size) training locations. The results are depicted in Figure 5. As visible in the figure, the usage of a hexagonal grid yields slightly better localization errors in contract to a squared grid, while the random grid yields the worst performance. Also in accordance to the statement given in Theorem 4.2, the increase in the number of training locations improves the performance of fingerprinting algorithms. Example-wise, an average accuracy improvement of roughly 30% is achieved in case the grid size is reduced from 3 m to 1 m.

Second, we evaluate the statement given in Theorem 4.1, concerning the number of observations at both training locations and in the runtime phase of fingerprinting. We stabilize the

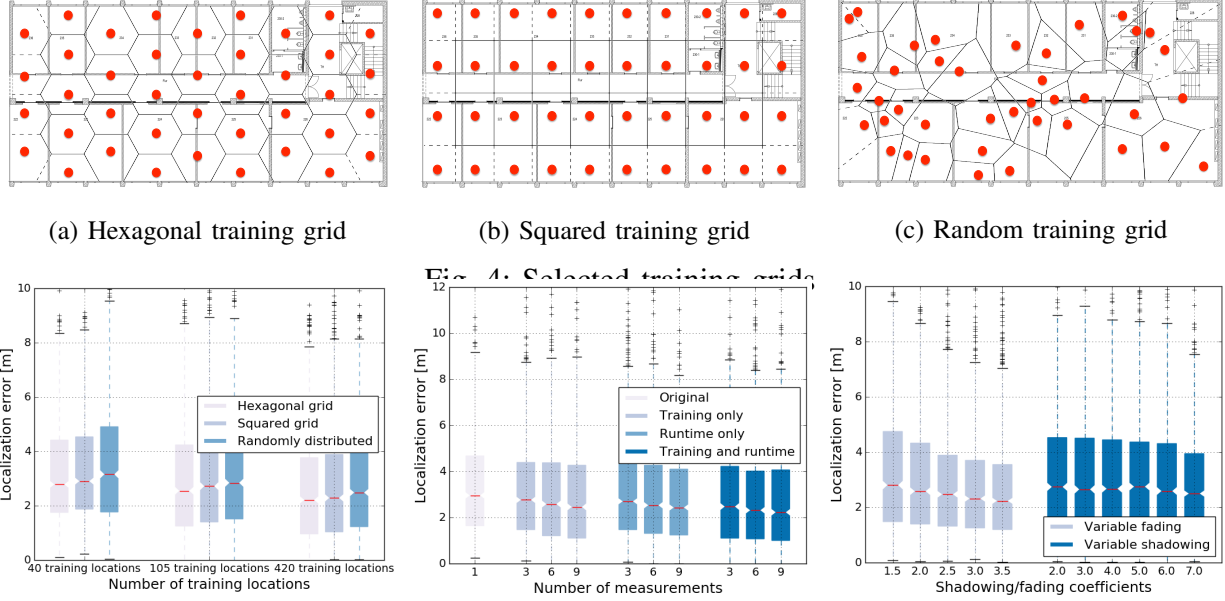


Fig. 5: Localization error vs. grid selection

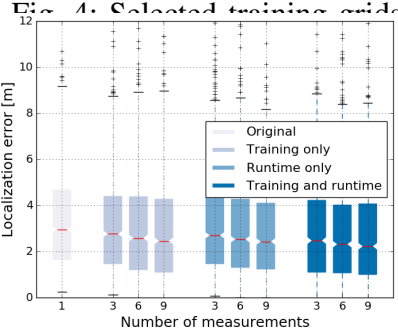


Fig. 6: Localization error vs. number of measurements

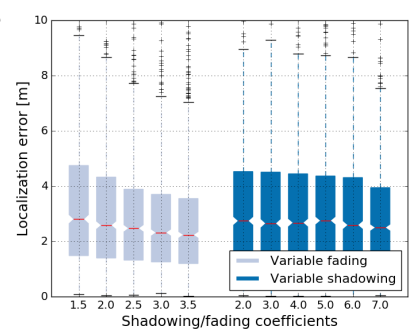


Fig. 7: Localization error vs. shadowing and fading

training grid to a hexagonal one of 105 training locations. Furthermore, we increase the number of observations from 1 to 15 with a step of 5 in both training and runtime phase of fingerprinting. The results are depicted in Figure 6. As visible in the figure, in comparison to the basic case where only one measurement is taken in both training and runtime phase, increase in the number of measurements generally reduces the localization error. Furthermore, the reduction of errors is higher in case one increases the number of measurements in both runtime and training phase, in contrast to increasing this number in one phase only. For example, the error is in average reduced by roughly 15% in case of the increase in the number of measurements from 1 to 9 in one phase only. For the same increase, but in both phases, the average error is reduced by more than 25%.

Third, we examine the statement given in Section V-C, stating that an increase in fading and shadowing propagation characteristics benefits the accuracy of fingerprinting. In our simulation model, fading is characterized by the path-loss exponent  $\gamma$ , while shadowing depends on the wall attenuation term  $l_w$ . Therefore, to evaluate the fading effect on the performance we increase the path-loss exponent from 1.5 to 3.5, while for characterizing the effect of shadowing we increase the wall attenuation term from 2 to 7 dBm. The achieved localization errors for such scenarios



are given in Figure 7. The results demonstrate that the increase in both fading and shadowing yields benefits for the performance of fingerprinting algorithms. For example, an increase in the path-loss exponent from 1.5 to 3.5 yields in average roughly 25% reduction in localization errors, while the increase in the wall attenuation term from 2 to 7 dBm results in roughly 20% decrease in localization errors.

Finally, we evaluate the statement given in Proposition 3 and discussion in Section V, concerning the number and locations of APs. We start from the basic scenario with 4 APs (AP 1, AP 2, AP 3, AP 4) depicted in Figure 8. We then introduce two additional APs (AP 5 and AP 6 in Figure 8) in the environmental, where their locations are selected either randomly or based on Voronoi vertices. Voronoi vertices-based selection places new APs at locations that are the farthest from the locations of existing APs. Furthermore, based on the locations of now 6 APs, we introduced additional 4 APs, both randomly and, as depicted in Figure 8, based on Voronoi vertices. The localization errors for such scenarios are depicted in Figure 9. As visible in the figure, the increase in the number of APs generally notably improves the performance of fingerprinting. Secondly, in comparison to 5 different random sections, both in case of 6 and 10 APs, Voronoi vertices-based selection generally yields more than 15% better performance. In other words, selection of APs in a way that they are the farthest from the locations of existing APs outperforms other selections. In conclusion, the results derived by simulation with a higher level of realism are consistent with the developed theory.

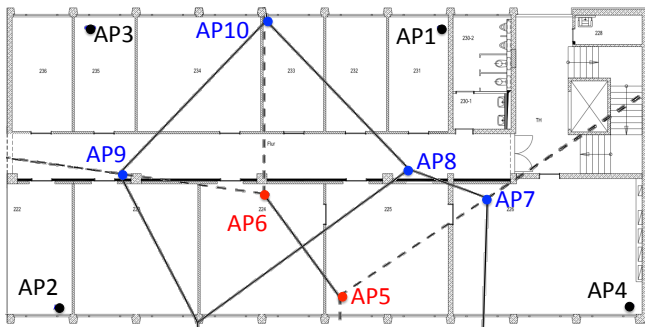


Fig. 8: Voronoi vertices-based selection of APs' locations

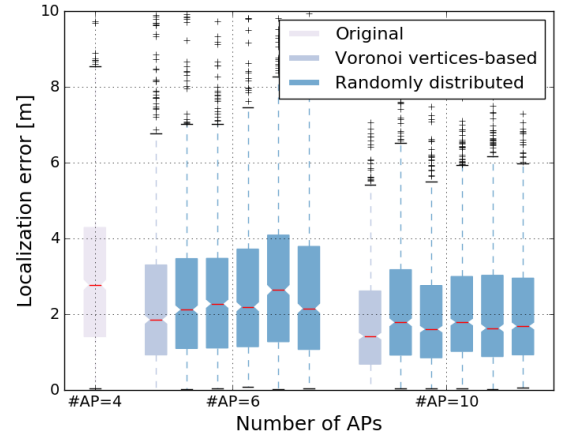


Fig. 9: Localization error vs. number and locations of APs

### B. Experimental Results

The TWIST testbed is specifically designed for indoor localization performance evaluation purposed experimentation. It features an automated experimentation without the need of a test-person, highly accurate ground-truth positioning, minimization and monitoring of external influences such as interference, and immediate calculation and storage of performance results [36] aligned with the EVARILOS Benchmarking Handbook (EBH) [37]. The EBP provides guidelines for objective evaluation of indoor localization solutions.

Training database has been created by collecting 20 RSS values from four APs in 41 training locations, as indicated in Figure 10a). The evaluation locations used in the evaluation are shown also in Figure 10, with their selection being based on the guidelines from the EBP.



Fig. 10: Training and evaluation locations

First we evaluate the claim that the main benefit of using multiple APs is the reduction of the far anchors effect. In other words, the localization errors increase with the increase in the distance between an AP and a target node. Figure 11 presents spatial distribution of errors for three different situation, each one using different number of APs for localization. As visible in Figure 11a, those locations farther away from AP 1 tend to have larger errors. This is because the inverse relation of RSS values with distance such that those points closer to AP1 have a finer RSS granularity. It is interesting to see that fingerprinting localization algorithm performs acceptable in indoor environment even with one anchor due to shadowing and multipath effects. Furthermore, the APs are added according to the guideline discussed in the paper. In the first step, AP 2 is deployed at the farthest location from AP 1. It can be seen in Figure 11b that such placement mainly decreases the errors at locations close to AP 2. Second, in Figure 11c, four APs are deployed at four corners of the testbed, which again improves the localization error.

Localization errors for different combinations of APs are given in Figure 12. As visible from the figure, an increase in the number of APs significantly reduces the error, e.g. in average the error is reduced by roughly 50% in case the number of APs increases from 2 to 4. The results derived in a realistic environment where the assumptions from the theoretical framework do not necessarily hold are in line with the developed theory.

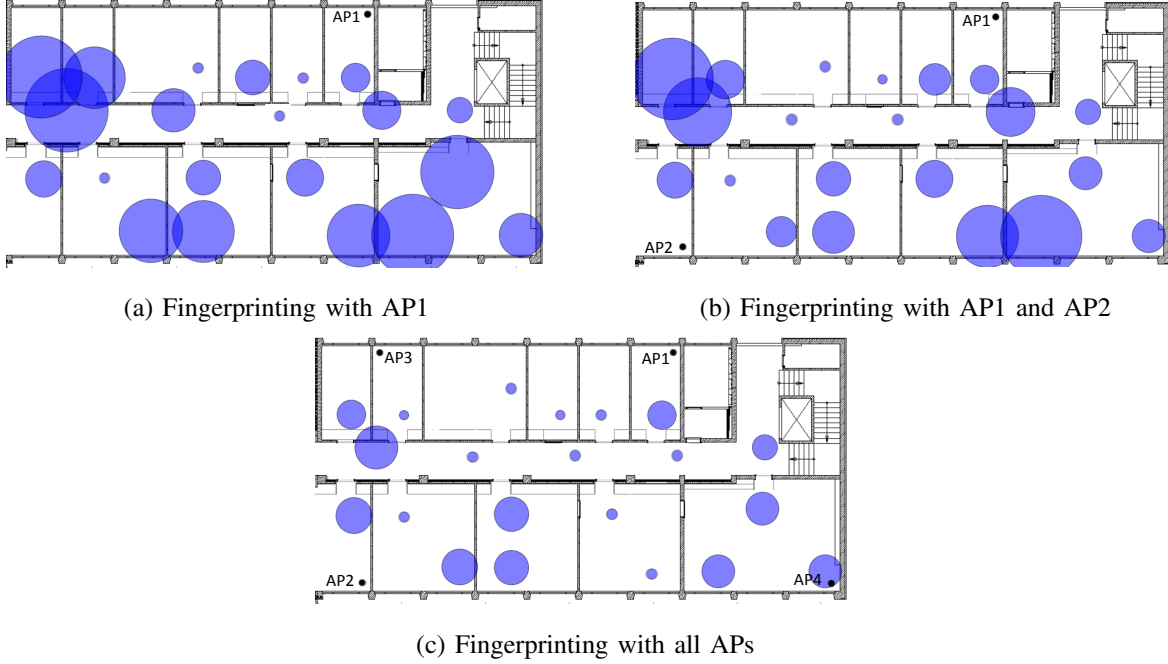


Fig. 11: Spatial distribution of errors with different AP selection

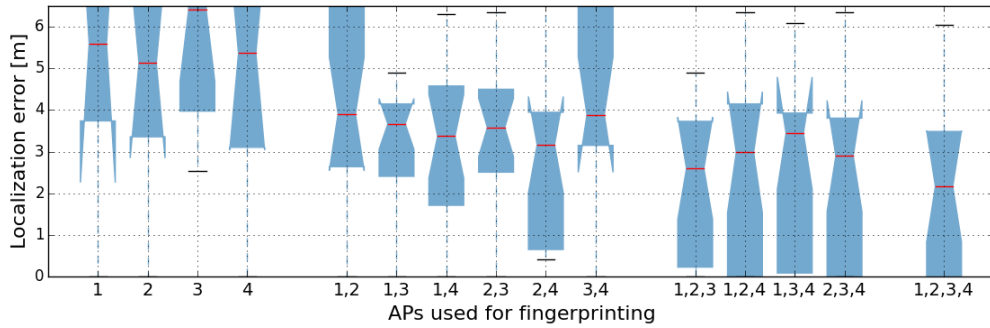


Fig. 12: Localization error vs. number of APs

## VII. CONCLUSION

Fingerprinting algorithms are generally multi-faceted and complex, hence it is hard to directly study the effect of their operational blocks and environmental parameters on their performance. In this work, we introduced a theoretical framework for analysis of fingerprinting algorithms

with the goal of providing guidelines on their design and performance analysis. We have shown that the performance of fingerprinting depends on the used fingerprinting feature and the dependence of this feature on the location in a fingerprinting space. We made a connection between fingerprinting and hypothesis testing problem. It was shown that the accuracy of fingerprinting algorithms is related to KL divergence between probability distribution of the selected feature at two locations. We additionally discussed the effect of the number of measurements at each location, as well as the effect of the training grid constellations. This framework was instantiated for RSS-based fingerprinting algorithms and guidelines were provided for anchor placement. It has been shown that shadowing and fading act in favor of fingerprinting algorithms by creating more variability in fingerprints of different locations. Simulations and experimental results demonstrate the consistency of the framework with realistic localization scenarios. The framework is promising for considering various research problems in fingerprinting scenarios.

## APPENDIX

### A. Proof of Proposition 3

Suppose that  $D(\mathbb{P}_{\mathbf{X}|\mathbf{u}_1} \parallel \mathbb{P}_{\mathbf{X}|\mathbf{u}_2}) \leq L$ . Using (11) for an anchor in origin, and the points at  $u_1 > 0$  and  $u_2 > 0$ , we have:

$$D(\mathbb{P}_{\mathbf{X}|\mathbf{u}_1} \parallel \mathbb{P}_{\mathbf{X}|\mathbf{u}_2}) = \frac{P_T^2}{2N_1} \left( \frac{1}{u_1^\alpha} - \frac{1}{u_2^\alpha} \right)^2.$$

A simple usage of mean value theorem implies that for some  $\beta \in [0, 1]$ :

$$\left| \frac{1}{u_1^\alpha} - \frac{1}{u_2^\alpha} \right| = |u_1 - u_2| \left| \frac{\alpha}{(\beta u_1 + (1 - \beta)u_2)^{\alpha+1}} \right|.$$

For this case if  $D(\mathbb{P}_{\mathbf{X}|\mathbf{u}_1} \parallel \mathbb{P}_{\mathbf{X}|\mathbf{u}_2}) = L$ , then:

$$|u_1 - u_2|^2 \leq \frac{L}{\kappa} (\beta u_1 + (1 - \beta)u_2)^{2\alpha+2}$$

where  $\kappa = \frac{\alpha^2 P_T^2}{2N_1}$ . Since the localization space is bounded inside  $[0, D]$  then  $\beta u_1 + (1 - \beta)u_2 \leq D$  and therefore:

$$|u_1 - u_2| \leq \frac{D^{\alpha+1}}{\alpha P_T} \sqrt{2N_1 L}.$$

### B. Proof of Theorem 4.1

Consider two points  $\mathbf{u}_1$  and  $\mathbf{u}_2$  in localization space. The theorem is equivalent to find bounds on missed detection and false alarm in a hypothesis testing scenario. The equivalent hypothesis testing problem is Stein's lemma and the proof is well known (for instance [24], [26]). Nonetheless the sketch of proof is provided here. For the known probabilities  $\mathbb{P}_{X|\mathbf{u}_1}$  and  $\mathbb{P}_{X|\mathbf{u}_2}$ , the *typical* set is defined as:

$$A_{\epsilon'}^n(X) = \left\{ \mathbf{x} \in \mathcal{X}^n : \left| \frac{1}{n} \sum_{i=1}^n \log \frac{\mathbb{P}_{X|\mathbf{u}_2}(x_i)}{\mathbb{P}_{X|\mathbf{u}_1}(x_i)} + D(\mathbb{P}_{X|\mathbf{u}_1} \parallel \mathbb{P}_{X|\mathbf{u}_2}) \right| \leq \epsilon' \right\}.$$

If the fingerprint is constructed at  $\mathbf{u}_1$ , the law of large numbers implies that the term  $\frac{1}{n} \sum_{i=1}^n \log \frac{\mathbb{P}_{X|\mathbf{u}_2}(x_i)}{\mathbb{P}_{X|\mathbf{u}_1}(x_i)}$  tends to its average value as  $n$  goes to infinity, which is  $-D(\mathbb{P}_{X|\mathbf{u}_1} \parallel \mathbb{P}_{X|\mathbf{u}_2})$ . Therefore for large enough  $n$ , it can be seen:

$$\alpha(\mathbf{u}_1, \mathbf{u}_2) = \mathbb{P}_{\mathbf{X}|\mathbf{u}_1}(\mathbf{X} \notin A_{\epsilon'}^n(X)) \leq \epsilon. \quad (15)$$

Indeed,  $A_{\epsilon'}^n(X)$  acts as a decision region for  $\mathbb{P}_{X|\mathbf{u}_1}$  and the preceding inequality guarantees that one can correctly identify  $\mathbf{u}_1$  based on the measured fingerprint  $\mathbf{X}$  with probability bigger than  $1 - \epsilon$ . To bound  $\beta(\mathbf{u}_1, \mathbf{u}_2)$ , suppose that the samples are obtained at  $\mathbf{u}_2$ .  $\beta(\mathbf{u}_1, \mathbf{u}_2)$  can be bounded from the above as follows:

$$\begin{aligned} \beta(\mathbf{u}_1, \mathbf{u}_2) &= \mathbb{P}_{\mathbf{X}|\mathbf{u}_2}(\mathbf{X} \in A_{\epsilon'}^n(X)) = \sum_{\mathbf{x} \in A_{\epsilon'}^n(X)} \mathbb{P}_{\mathbf{X}|\mathbf{u}_2}(\mathbf{x}) \leq \\ &\sum_{\mathbf{x} \in A_{\epsilon'}^n(X)} \mathbb{P}_{\mathbf{X}|\mathbf{u}_1}(\mathbf{x}) \exp(-n(D(\mathbb{P}_{X|\mathbf{u}_1} \parallel \mathbb{P}_{X|\mathbf{u}_2}) - \epsilon')) \leq \exp(-n(D(\mathbb{P}_{X|\mathbf{u}_1} \parallel \mathbb{P}_{X|\mathbf{u}_2}) - \epsilon')). \end{aligned} \quad (16)$$

On the other hand, the inner bound is obtained as follows:

$$\begin{aligned} \beta(\mathbf{u}_1, \mathbf{u}_2) &= \mathbb{P}_{\mathbf{X}|\mathbf{u}_2}(\mathbf{X} \in A_{\epsilon'}^n(X)) = \sum_{\mathbf{x} \in A_{\epsilon'}^n(X)} \mathbb{P}_{\mathbf{X}|\mathbf{u}_2}(\mathbf{x}) \geq \\ &\sum_{\mathbf{x} \in A_{\epsilon'}^n(X)} \mathbb{P}_{\mathbf{X}|\mathbf{u}_1}(\mathbf{x}) \exp(-n(D(\mathbb{P}_{X|\mathbf{u}_1} \parallel \mathbb{P}_{X|\mathbf{u}_2}) + \epsilon')) = \alpha(\mathbf{u}_1, \mathbf{u}_2) \exp(-n(D(\mathbb{P}_{X|\mathbf{u}_1} \parallel \mathbb{P}_{X|\mathbf{u}_2}) + \epsilon')) \\ &\geq (1 - \epsilon) \exp(-n(D(\mathbb{P}_{X|\mathbf{u}_1} \parallel \mathbb{P}_{X|\mathbf{u}_2}) + \epsilon')). \end{aligned} \quad (17)$$

By taking the logarithm from both sides and tending  $n$  to infinity, the theorem is proved.

### C. Proof of Theorem 4.2

Consider the set  $C_n$ , called the critical region, defined as follows:

$$C_n = \left\{ \mathbf{x} : \inf_{\mathbf{u}_1 \in \hat{\mathcal{V}}_{\mathbf{v}_1}} D(\mathbb{Q}_{X|\mathbf{u}} \| \mathbb{P}_{X|\mathbf{u}_1}) \geq \delta_n \right\}$$

where  $\mathbb{Q}_{X|\mathbf{u}}$  is the empirical distribution of  $\mathbf{x}$  measured at  $\mathbf{u}$  and  $\delta_n = \Omega(\frac{\log n}{n})$ . Then using Theorem 2.3 in [38], it can be seen that if  $\mathbf{X}$  follows the distribution  $\mathbb{Q}_{X|\mathbf{u}}$  for  $\mathbf{u} \in \hat{\mathcal{V}}_{\mathbf{v}_1}$ , then  $\Pr(\mathbf{X} \in C_n) \leq \epsilon$ . Using the same theorem, the second error is shown to decay exponentially with the exponent indicated in the theorem. So  $(C_n)^s$  can be a decision region for finding the closest fingerprint.

### REFERENCES

- [1] C. Medina, J. Segura, and . De la Torre, "Ultrasound Indoor Positioning System Based on a Low-Power Wireless Sensor Network Providing Sub-Centimeter Accuracy," *Sensors*, vol. 13, no. 3, pp. 3501–3526, Mar. 2013.
- [2] E. Brassart, C. Pegard, and M. Mouaddib, "Localization using infrared beacons," *Robotica*, vol. 18, no. 02, pp. 153–161, Mar. 2000.
- [3] M. Erol-Kantarci, H. T. Mouftah, and S. Oktug, "A Survey of Architectures and Localization Techniques for Underwater Acoustic Sensor Networks," *IEEE Communications Surveys Tutorials*, vol. 13, no. 3, pp. 487–502, 2011.
- [4] I. Amundson and X. D. Koutsoukos, "A Survey on Localization for Mobile Wireless Sensor Networks," in *Mobile Entity Localization and Tracking in GPS-less Environments*, ser. Lecture Notes in Computer Science, R. Fuller and X. D. Koutsoukos, Eds. Springer Berlin Heidelberg, 2009, no. 5801, pp. 235–254, doi: 10.1007/978-3-642-04385-7\_16.
- [5] F. Seco *et al.*, "A survey of mathematical methods for indoor localization," in *Intelligent Signal Processing*, 2009.
- [6] D. Milioris, G. Tzagkarakis, A. Papakonstantinou, M. Papadopouli, and P. Tsakalides, "Low-dimensional signal-strength fingerprint-based positioning in wireless {LANs}," *Ad Hoc Networks*, vol. 12, pp. 100 – 114, 2014.
- [7] V. Honkavirta *et al.*, "A Comparative Survey of WLAN Location Fingerprinting Methods," in *WPNC 2009*. IEEE, 2009, pp. 243–251.
- [8] D. Milioris *et al.*, "Low-Dimensional Signal-Strength Fingerprint-based Positioning in Wireless LANs," *Ad Hoc Networks*, 2011.
- [9] C. Laoudias, P. Kemppi, and C. Panayiotou, "Localization using radial basis function networks and signal strength fingerprints in wlan," in *Global Telecommunications Conference, 2009. GLOBECOM 2009. IEEE*, 2009, pp. 1–6.
- [10] S. Bai and T. Wu, "Analysis of k-means algorithm on fingerprint based indoor localization system," in *Microwave, Antenna, Propagation and EMC Technologies for Wireless Communications*, 2013.
- [11] C. Steiner and A. Wittneben, "Efficient training phase for ultrawideband-based location fingerprinting systems," *Signal Processing, IEEE Transactions on*, vol. 59, no. 12, pp. 6021–6032, 2011.
- [12] J. Machaj, P. Brida, and B. Tatarova, "Impact of the number of access points in indoor fingerprinting localization," in *Radioelektronika*, 2010, pp. 1–4.
- [13] T.-N. Lin *et al.*, "Performance comparison of indoor positioning techniques based on location fingerprinting in wireless networks," in *Wireless Networks, Communications and Mobile Computing*, vol. 2, 2005, pp. 1569–1574.

- [14] K. Kaemarungsi and P. Krishnamurthy, "Modeling of indoor positioning systems based on location fingerprinting," in *INFOCOM*, 2004, pp. 1012–1022.
- [15] K. Kaemarungsi, "Efficient design of indoor positioning systems based on location fingerprinting," in *Wireless Networks, Communications and Mobile Computing*, vol. 1, 2005, pp. 181–186.
- [16] Y. Wen, X. Tian, X. Wang, and S. Lu, "Fundamental limits of RSS fingerprinting based indoor localization," in *2015 IEEE Conference on Computer Communications (INFOCOM)*, Apr. 2015, pp. 2479–2487.
- [17] G. Ding *et al.*, "Overview of received signal strength based fingerprinting localization in indoor wireless lan environments," in *Microwave, Antenna, Propagation and EMC Technologies for Wireless Communications*, 2013.
- [18] W. Meng *et al.*, "Secure and robust wi-fi fingerprinting indoor localization," in *Indoor Positioning and Indoor Navigation*, 2011, pp. 1–7.
- [19] A. Mahtab Hossain, Y. Jin, W.-S. Soh, and H. N. Van, "Ssd: A robust rf location fingerprint addressing mobile devices' heterogeneity," *Mobile Computing, IEEE Transactions on*, vol. 12, no. 1, pp. 65–77, 2013.
- [20] C. Beder and M. Klepal, "Fingerprinting based localisation revisited: A rigorous approach for comparing rssi measurements coping with missed access points and differing antenna attenuations," in *Indoor Positioning and Indoor Navigation*, 2012, pp. 1–7.
- [21] P. M. Scholl, S. Kohlbrecher, V. Sachidananda, and K. Van Laerhoven, "Fast indoor radio-map building for RSSI-based localization systems." *IEEE*, Jun. 2012, pp. 1–2.
- [22] F. Lemic, A. Behboodi, V. Handziski, and A. Wolisz, "Experimental Decomposition of the Performance of Fingerprinting-based Localization Algorithms," in *5th International Conference on Indoor Positioning and Indoor Navigation*, 2014.
- [23] E. L. Lehmann and J. P. Romano, *Testing Statistical Hypotheses*. Springer Science & Business Media, Mar. 2006.
- [24] A. Dembo and O. Zeitouni, *Large Deviations Techniques and Applications*, ser. Stochastic Modelling and Applied Probability. Berlin, Heidelberg: Springer Berlin Heidelberg, 2010, vol. 38. [Online]. Available: <http://link.springer.com/10.1007/978-3-642-03311-7>
- [25] W. Hoeffding, "Probability Inequalities for Sums of Bounded Random Variables," *Journal of the American Statistical Association*, vol. 58, no. 301, pp. 13–30, Mar. 1963.
- [26] I. Csiszar and J. Korner, *Information Theory: Coding Theorems for Discrete Memoryless Systems*. Orlando, FL, USA: Academic Press, Inc., 1982.
- [27] U. Erez *et al.*, "Lattices which are good for (almost) everything," *Information Theory, IEEE Transactions on*, vol. 51, no. 10, pp. 3401–3416, 2005.
- [28] D. Tse and P. Viswanath, *Fundamentals of wireless communication*. Cambridge, UK ; New York: Cambridge University Press, 2005, oCLC: ocm57751753.
- [29] F. Lemic, "Benchmarking of Quantile based Indoor Fingerprinting Algorithm," Tech. Rep. TKN-14-001, 2014.
- [30] F. Lemic, V. Handziski, A. Wolisz *et al.*, "Experimental Evaluation of RF-based Indoor Localization Algorithms Under RF Interference," in *5th International Conference on Localization and GNSS*, 2015.
- [31] A. Borrelli, C. Monti, M. Vari, and F. Maz, "Channel models for ieee 802.11 b indoor system design," in *ICC'04*, vol. 6. IEEE, 2004, pp. 3701–3705.
- [32] G. Caso *et al.*, "On the Applicability of Multi Wall Multi Floor Propagation Models to WiFi Fingerprinting Indoor Positioning," in *FABULOUS'15*, 2015.
- [33] F. Lemic, V. Handziski, G. Caso *et al.*, "Towards The Extrapolation of WiFi RSSI-based Fingerprinting Performance Across Environments," in *Proc. of the 17th Workshop on Mobile Computing Systems and Applications (ACM HotMobile'16)*, 2016.

- [34] F. Lemic, J. Martin, C. Yarp, D. Chan, V. Handziski, R. Brodersen, G. Fettweis, A. Wolisz, and J. Wawrzyniek, "Localization as a Feature of mmWave Communication," in *IEEE Wireless Communications and Mobile Computing*, 2016.
- [35] F. Lemic, J. Büsch, M. Chwalisz, V. Handziski, and A. Wolisz, "Infrastructure for Benchmarking RF-based Indoor Localization under Controlled Interference," in *Ubiquitous Positioning, Navigation and Location-Based Services*, 2014.
- [36] F. Lemic *et al.*, "Web-based Platform for Evaluation of RF-based Indoor Localization Algorithms," in *Communications Workshops (ICC), 2015 IEEE International Conference on Communications*. IEEE, 2015.
- [37] T. Van Haute *et al.*, "The EVARILOS Benchmarking Handbook: Evaluation of RF-based Indoor Localization Solutions," in *Measurement-based Experimental Research, Methodology and Tools (MERMAT'13)*, 2013.
- [38] I. Csiszr and P. C. Shields, "Information Theory and Statistics: A Tutorial," *Foundations and Trends in Communications and Information Theory*, vol. 1, no. 4, pp. 417–528, 2004.